

Estudio de correlación de las variables significativas analizadas en el censo económico 2010

Correlation Study of the significant variables analyzed in the 2010 Economic Census

*Raúl Álvarez Gual**

Resumen

Este artículo presenta un estudio del análisis de correlación de las variables más significativas presentadas por el Instituto Nacional de Estadísticas y Censo (INEC) del Censo económico 2010. El estudio hace un enfoque utilizando las correlaciones lineales de Pearson, y se pretende eliminar las correlaciones sin sentido encontradas en este análisis por medio de correlaciones parciales.

Palabras clave

Econometría, correlación, covarianza, actividades económicas.

Abstract

This article presents a study of correlation analysis about the most significant variables presented by the National Institute of Statistics and Census of economical census in 2010. The study makes focus using the linear correlation of Pearson and it is pretended to eliminate nonsense correlations found in this analysis by the use partial correlations.

Keywords

Econometrics, correlation, covariance, economic activities.

Forma sugerida de citar: Álvarez, Raúl. 2011. Estudio de correlación de las variables significativas analizadas en el censo económico 2010. Retos 2. Julio/diciembre. Pp. 105-124.

* Docente UPS. Ingeniero en Estadística e Informática. Director del Centro de Investigaciones Económicas y Empresariales (CIEE) de la Universidad Politécnica Salesiana. E mail: ralvarezg@ups.edu.ec.

Introducción

La Econometría es el campo de la economía que se ocupa de la medición empírica tanto sean éstas como la estimación, la inferencia y las predicciones de variables; la econometría también analiza las relaciones entre variables que establece la teoría económica, a través de la aplicación de métodos estadísticos, matemáticos y computacionales. El propósito fundamental es proporcionar contenido empírico a las relaciones teóricas. (Cisepa, 2003).

Este trabajo se centra en utilizar una técnica econométrica para determinar las relaciones existentes entre las siguientes variables: *Cantidad de personal* que trabajan en las empresas económicamente activas, los *Gastos por remuneración* que tienen las empresas, el *Financiamiento obtenido* en todo el período de estudio, los *Gastos por electricidad* generados, los *Ingresos netos obtenidos*, y el *uso de Internet*.

La herramienta para determinar la relación existente entre cada una de las variables es la el coeficiente de correlación poblacional ρ de Pearson, que se define como:

$$\rho = \frac{cov(x,y)}{\sqrt{\{var(x)var(y)\}}} = \frac{cov(x,y)}{\sigma_x \sigma_y} \quad (1)$$

Fuente: (Damodar Gujarati, 2010).

Donde la covarianza se calcula por medio de la fórmula:

$$Cov(x,y) = \frac{\sum x_i y_i}{n} - \bar{x} \bar{y} \quad (2)$$

Fuente: (Damodar Gujarati, 2010).

$$\bar{x} = \frac{\sum x_i}{n} \quad (3)$$

$$\bar{y} = \frac{\sum y_i}{n} \quad (4)$$

$$s_y = \sqrt{\frac{\sum (\bar{y} - y_i)^2}{n}} \quad y \quad s_x = \sqrt{\frac{\sum (\bar{x} - x_i)^2}{n}} \quad (5)$$

Fuente: (Damodar Gujarati, 2010).

El *coeficiente de correlación* es uno de los muchos coeficientes estadísticos para medir la relación lineal de varias variables, puede tomar valores entre -1 y 1, siendo su signo igual al de la covarianza ($cov(x,y)$).¹ Las correlaciones con valor absoluto 1 implican que existe una asociación matemática lineal perfecta, positiva o negativa, entre las dos variables y correlaciones iguales a 0 implican ausencia de asociación. Obviamente, las variables independientes tienen correlación 0, pero nuevamente, la independencia es condición suficiente pero no necesaria (Madrid).

Metodología

Los datos presentados por el Instituto Ecuatoriano de Estadísticas y Censos (INCEC) sobre el censo económico 2010 son muy enriquecedores para la realización de estudios econométricos. La metodología a utilizar en esta investigación la podemos resumir en cuatro partes:

Parte 1: Clasificación de variables y obtención de datos

- Identificación de las variables de investigación realizadas por el INEC.
- Escoger seis variables significativas de estudio y renombrarlas (Debido a que los nombres son muy amplios se acortarán los mismos).

- c) Identificación de de las actividades económicas estudiadas por el INEC.
- d) Obtener una tabla de los valores totales por país por cada una de las variables escogidas.
- e) Obtener una tabla de los valores totales porcentuales por país por cada una de las variables escogidas.

Parte 2: Matriz de correlaciones

- a) Obtener una matriz de correlaciones de Pearson de las 6 variables de estudio.
- b) Realizar un diagrama de Iconografía de correlaciones para cada valor mayor a 0,5.

Parte 3: Correlaciones parciales y semiparciales

- a) Obtener una matriz de correlaciones de semiparciales combinadas para cada variable estática (fila de la matriz) y las otra par de variables como variables dinámicas (columna de la matriz).
- b) Realizar un diagrama de Iconografía de correlaciones de semiparciales para cada dato de cada par de variables dinámicas (columna de la matriz) en la que cada valor de su correlación parcial superaron a 0,5.

Parte 4: Análisis de las actividades económicas

- a) Obtener una matriz de correlaciones en donde cada actividad represente una variable y por cada columna agregar un 1 y el resto cero para que exista congruencia de valores entre variables y datos.
- b) Realizar un diagrama de iconografía de correlaciones para cada dato de cada par de variables en la que cada valor de su correlación parcial superaron a 0,3.

El estudio está orientado a demostrar la existencia de relaciones de dos variables que se encuentren involucradas en relaciones multivariadas.

Parte1.

Datos a utilizar

Los datos a utilizarse serán los proporcionados por el INEC y publicados en su página web.

Censo Económico 2010

El INEC entrega al país los resultados del Censo Nacional Económico (CENEC), realizado durante los meses de septiembre, octubre y noviembre de 2010, información de alta trascendencia con la que el Ecuador no contaba desde hace 30 años”.²

Se investigaron 511.130 establecimientos en áreas urbanas que completaran los 2.000 habitantes o más, también se consideraron cabeceras cantonales que no cumplieran con la cifra mencionada, e inclusive se tomaron en cuenta los corredores de vías principales consideradas como zonas económicas especiales y las grandes empresas.

La información recabada en el empadronamiento corresponde a la identificación y ubicación del establecimiento; clasificación de la actividad económica; ingresos, costos, gastos, existencias y activos fijos totales; personal ocupado y gastos en remuneraciones; financiamiento del establecimiento e información adicional del establecimiento.³

Variables de investigación:

Las variables que se investigaron en el Censo Económico 2010, son las siguientes:

Variable estudiada	
1	Nombre y Razón Social del establecimiento.
2	Dirección del establecimiento
3	Año de constitución del establecimiento.
4	Calificación artesanal.
5	Tenencia del establecimiento propio o arrendado.
6	Tipo de establecimiento (único, matriz, sucursal y auxiliar).
7	RUC del establecimiento.
8	Personal ocupado y gasto en remuneraciones.
9	Actividad principal y secundaria del establecimiento.
10	Principal bien o servicio.
11	Materia prima.
12	Fecha de inicio de la actividad principal.
13	Comercio al por mayor y al por menor.
14	Principal cliente.
15	Costos, gastos, ingresos, existencia y activos fijos.
16	Forma del establecimiento (jurídica y no jurídica).
17	Fuentes de financiamiento.
18	Requerimiento de financiamiento.
19	Investigaciones o estudios de mercado.
20	Gasto en consumo de energía y combustibles.
21	Manejo de desechos y/o remediación ambiental.
22	Actividades de investigación y desarrollo.
23	Gastos en capacitación y formación.
24	Uso de internet.
25	Afiliación a gremios, cámaras u otras formas de asociación empresarial
26	Nombre y razón social de la matriz
27	Dirección de la matriz.
28	Nombre del gerente y propietario.
29	Datos del informante.
30	Listado de sucursales.

Fuente: INEC-Censo Económico 2010

Elaboración: El Autor

Para este análisis de correlaciones bivariadas de variables significativas se ha escogido un total de seis por todo el

país, es decir, todas y cada una de las variables que se muestran en la siguiente tabla, son valores nacionales.

No	Variable estudiada	Variable estudiada resumida ⁴
1	Personal ocupado, según clasificación CIIU 4.0 de la actividad económica principal a nivel nacional, 2009	Personal
2	Gastos en remuneraciones totales de los establecimientos únicos y matrices censados en el 2009, según clasificación CIIU 4.0 de la actividad principal	Gastos remuneración
3	Monto de financiamiento obtenido en el 2009 de establecimientos económicos únicos y matrices censados, según CIIU 4.0 de la actividad principal nacional.	Financiamiento
4	Ingresos anuales de los establecimientos.	Ingresos
5	Gastos de energía eléctrica anual de los establecimientos únicos y matrices censados en el 2009; según CIIU de la actividad económica principal a nivel nacional.	Gastos de energía
6	Uso del Internet en establecimientos únicos y matrices censados, según clasificación CIIU 4.0 de la actividad principal a nivel nacional, 2009.	Uso de Internet

Fuente: INEC-Censo Económico 2010

Elaboración: El Autor

Actividades económicas

El INEC se ha basado en la clasificación de las actividades económicas por medio de la Clasificación Industrial Internacional Uniforme de las Actividades Económicas (CIIU) emitido por del departamento de Asuntos Económicos y Sociales de la

división Estadística de las Naciones Unidas.⁵

Esta clasificación hace una categorización de las actividades económicas productivas para que se puedan utilizar en la elaboración de estudios estadísticos. En la Tabla 1. se muestra los nombres generales de cada una de las actividades:

Tabla 1. Actividades Económicas según CIIU 4.0

	Actividades Económicas	Nombre Actividades Económicas resumida para caso de estudio
1	Agricultura, ganadería, silvicultura y pesca.	A1
2	Explotación de minas y canteras.	A2
3	Industrias manufactureras.	A3
4	Suministro de electricidad, gas, vapor y aire acondicionado.	A4
5	Distribución de agua; alcantarillado, gestión de desechos y actividades de saneamiento.	A5
6	Construcción.	A6
7	Comercio al por mayor y al por menor; reparación de vehículos automotores y motocicletas.	A7
8	Transporte y almacenamiento.	A8
9	Actividades de alojamiento y de servicio de comidas.	A9

10	Información y comunicación.	A10
11	Actividades financieras y de seguros.	A11
12	Actividades inmobiliarias.	A12
13	Actividades profesionales, científicas y técnicas.	A13
14	Actividades de servicios administrativos y de apoyo.	A14
15	Administración pública y defensa; planes de seguridad social de afiliación obligatoria.	A15
16	Enseñanza.	A16
17	Actividades de atención de la salud humana y de asistencia social.	A17
18	Artes, entretenimiento y recreación.	A18
19	Otras actividades de servicios.	A19
20	Actividades de organizaciones y órganos extraterritoriales.	A20

Fuente: Archivo Clasificación de actividad CIIU 4.0.pdf proporcionado por el INEC⁷ y nomenclatura.

Elaboración: El Autor

Datos del Censo Económico 2010

En los datos proporcionado por el INEC en su archivo *Tabulados Censo Nacional Económico 2009.xlsx*, se muestra los valores tabulados y resumidos por cada una de las variables investigadas clasificadas por cantones, provincias, por regiones. Para nuestro caso de estudio de las seis variables se muestra el Total por país, se lo describe de la siguiente manera:

Personal. Es el personal económicamente activo con o sin remuneración que prestó sus servicios en el año 2009 en el establecimiento económico censado.

Gastos remuneración. Es el total de gastos por remuneración por sucursales y

matrices que invirtió en el año 2009 cada establecimiento económico censado.

Financiamiento. Es el total del monto de financiamiento por sucursales y matrices que obtuvieron en el año 2009 cada establecimiento económico censado.

Gastos electricidad. Es el total de gastos de electricidad por sucursales y matrices que invirtieron en el año 2009 cada establecimiento económico censado.

Ingresos. Es total de ingresos que obtuvieron en el año 2009 en el establecimiento económico censado.

Internet. Es el número de establecimientos con servicio de la Internet que poseían en el año 2009 en el establecimiento económico censado.

Tabla 2. Valores de totales de las variables analizadas en el Censo Económico 2010

Actividades económicas		Personal	Gastos remuneración	Financiamiento	Uso de Internet	Gastos de electricidad	Ingresos
1	Agricultura, ganadería, silvicultura y pesca.	37,280	201,762,316	217,254,696	233	9,715,052	1,111,027,695.29
2	Explotación de minas y canteras.	13,795	1,550,179,447	188,502,287	64	1,349,329	3,300,091,307.43
3	Industrias manufactureras.	266,908	1,559,031,898	908,325,758	5,003	292,552,934	40,509,030,141.86
4	Suministro de electricidad, gas, vapor y aire acondicionado.	8,692	84,669,334	38,653,925	55	2,245,059	631,069,203.00
5	Distribución de agua; alcantarillado, gestión de desechos y actividades de saneamiento.	6,354	68,770,619	7,117,614	59	2,496,620	440,926,822.86
6	Construcción.	32,849	231,752,228	125,511,457	666	4,586,981	2,029,547,442.43
7	Comercio al por mayor y al por menor; reparación de vehículos automotores y motocicletas.	611,390	2,211,339,338	1,271,503,883	15,577	118,401,517	60,314,136,470.86
8	Transporte y almacenamiento.	49,200	393,180,818	158,516,447	1,173	21,734,024	3,869,350,477.43
9	Actividades de alojamiento y de servicio de comidas.	153,295	278,091,408	101,948,601	1,974	32,725,490	2,779,681,688.86
10	Información y comunicación.	51,841	222,745,579	534,384,153	6,248	14,318,533	3,014,685,793.43
11	Actividades financieras y de seguros.	48,145	667,558,917	1,098,648,713	883	30,068,989	7,131,480,037.71
12	Actividades inmobiliarias.	9,701	204,162,153	54,774,759	587	6,016,647	840,151,838.29
13	Actividades profesionales, científicas y técnicas.	62,318	595,415,119	320,024,070	5,245	12,560,049	3,319,334,865.71
14	Actividades de servicios administrativos y de apoyo.	63,399	510,020,112	45,268,402	1,964	4,999,509	1,665,855,251.71
15	Administración pública y defensa; planes de seguridad social de afiliación obligatoria.	185,042	1,805,854,708	667,404,688	874	21,970,094	6,012,423,998.86
16	Enseñanza.	227,688	1,790,740,284	214,542,476	3,801	17,927,053	4,398,741,476.29
17	Actividades de atención de la salud humana y de asistencia social.	106,634	848,608,672	72,006,779	3,882	14,677,192	2,519,185,031.43

Actividades económicas		Personal	Gastos remuneración	Financiamiento	Uso de Internet	Gastos de electricidad	Ingresos
18	Artes, entretenimiento y recreación.	22,676	272,540,577	7,668,318	517	4,167,498	725,521,550.86
19	Otras actividades de servicios.	102,072	227,501,250	53,940,180	2,569	14,192,311	1,248,154,150.71
20	Actividades de Organizaciones y Órganos Extraterri- toriales.	225	2,010,732	854,198	15	19,284	4,628,386.00
TOTALES:		2,059,504	13,725,935,509	6,086,851,404	51,389	626,724,165	145,865,023,631

Fuente: Tabulados Censo Económico 2009.xlsx[7] del sitio web <<http://www.inec.gov.ec/home/>>
Elaboración: El Autor

En vista de que los valores son muy altos en algunas de las variables y significativamente bajos en otras, es recomendable tratar el estudio por medio de ponderaciones porcentuales. Dividiendo cada valor con su respectivo total tenemos:

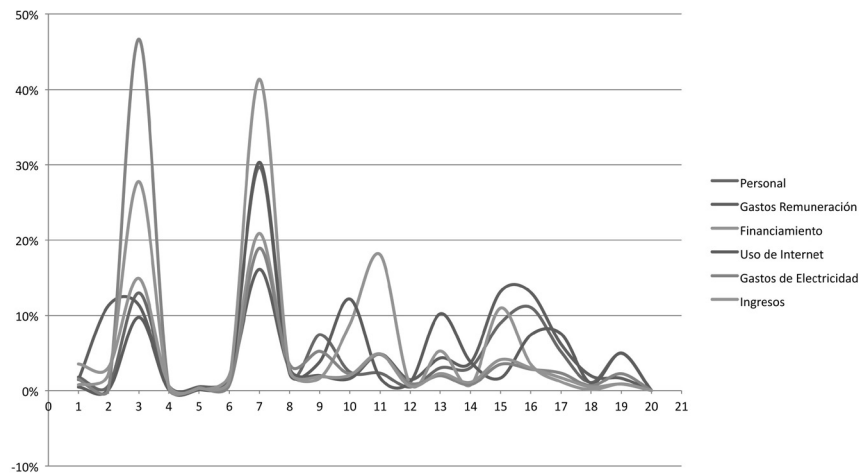
Tabla 3. Valores porcentuales de totales de las variables analizadas en el Censo Económico 2010

Actividades económicas		Personal	Gastos remuneración	Financiamiento	Uso de Internet	Gastos de electricidad	Ingresos
1	Agricultura, ganadería, silvicultura y pesca.	1.810%	1,470%	3,569%	0,453%	1,550%	0,762%
2	Explotación de minas y canteras.	0,670%	11,294%	3,097%	0,125%	0,215%	2,262%
3	Industrias manufactureras.	12.960%	11,358%	14,923%	9,736%	46,680%	27,772%
4	Suministro de electricidad, gas, vapor y aire acondicionado.	0,422%	0,617%	0,635%	0,107%	0,358%	0,433%
5	Distribución de agua; alcantarillado, gestión de desechos y actividades de saneamiento.	0,309%	0,501%	0,117%	0,115%	0,398%	0,302%
6	Construcción.	1,595%	1,688%	2,062%	1,296%	0,732%	1,391%
7	Comercio al por mayor y al por menor; reparación de vehículos automotores y motocicletas.	29,686%	16,111%	20,889%	30,312%	18,892%	41,349%
8	Transporte y almacenamiento.	2,389%	2,865%	2,604%	2,283%	3,468%	2,653%

	Actividades económicas	Personal	Gastos remuneración	Financiamiento	Uso de Internet	Gastos de electricidad	Ingresos
9	Actividades de alojamiento y de servicio de comidas.	7,443%	2,026%	1,675%	3,841%	5,222%	1,906%
10	Información y comunicación.	2,517%	1,623%	8,779%	12,158%	2,285%	2,067%
11	Actividades financieras y de seguros.	2,338%	4,863%	18,050%	1,718%	4,798%	4,889%
12	Actividades inmobiliarias.	0,471%	1,487%	0,900%	1,142%	0,960%	0,576%
13	Actividades profesionales, científicas y técnicas.	3,026%	4,338%	5,258%	10,206%	2,004%	2,276%
14	Actividades de servicios administrativos y de apoyo.	3,078%	3,716%	0,744%	3,822%	0,798%	1,142%
15	Administración pública y defensa; planes de seguridad social de afiliación obligatoria.	8,985%	13,157%	10,965%	1,701%	3,506%	4,122%
16	Enseñanza.	11,055%	13,046%	3,525%	7,397%	2,860%	3,016%
17	Actividades de atención de la salud humana y de asistencia social.	5,178%	6,183%	1,183%	7,554%	2,342%	1,727%
18	Artes, entretenimiento y recreación.	1,101%	1,986%	0,126%	1,006%	0,665%	0,497%
19	Otras actividades de servicios.	4,956%	1,657%	0,886%	4,999%	2,265%	0,856%
20	Actividades de organizaciones y órganos Extraterritoriales.	0,011%	0,015%	0,014%	0,029%	0,003%	0,003%
		100%	100%	100%	100%	100%	100%

Fuente: Tabulados Censo Económico 2003.xlsx del sitio web <<http://www.inec.gov.ec/home/>>
Elaboración: El autor.

Imagen 1. Porcentajes de totales PAÍS de las actividades económicas y variables de estudio.



Fuente: INEC- Censo Económico 2010
Elaboración: El Autor

Parte 2. Resultados

De acuerdo a la metodología propuesta se procede a obtener los coeficientes de correlación de Pearson entre cada una de las variables elegidas y por rama de actividad para determinar su relación existente. Se procede a obtener el coeficiente de correlación para

cada par de variables. Como ejemplo se muestra el Anexo 1.

El lector podrá observar que habría que realizar cada posible combinación por cada par de variables. Para evitar este cálculo matemático se muestra el resumen de la tabla de correlaciones obtenidas por el programa estadístico SPSS.

Tabla 4. Correlación de Pearson obtenida por SPSS.

		Personal	Gastos	Financiamiento	Internet	Energía	Ingresos
Personal	Pearson Correlation	1	0,766**	0,693**	0,868**	0,597**	0,897**
	Sig. (2-tailed)		0,000	0,001	0,000	0,005	0,000
	N	20	20	20	20	20	20
Gastos	Pearson Correlation	0,766**	1	0,641**	0,564**	0,496*	0,672**
	Sig. (2-tailed)	0,000		0,002	0,010	0,026	0,001
	N	20	20	20	20	20	20

		Personal	Gastos	Financiamiento	Internet	Energía	Ingresos
Financiamiento	Pearson Correlation	0,693**	0,641**		1	0,646**	0,617**
	Sig. (2-tailed)	0,001	0,002			0,002	0,004
	N	20	20	20	20	20	20
Internet	Pearson Correlation	0,868**	0,564**	0,646**		1	0,473*
	Sig. (2-tailed)	0,000	0,010	0,002			0,035
	N	20	20	20	20	20	20
Energía	Pearson Correlation	0,597**	0,496*	0,617**	0,473*		1
	Sig. (2-tailed)	0,005	0,026	0,004	0,035		
	N	20	20	20	20	20	20
Ingresos	Pearson Correlation	0,897**	0,672**	0,781**	0,821**	0,791**	
	Sig. (2-tailed)	0,000	0,001	0,000	0,000	0,000	
	N	20	20	20	20	20	20

** . Correlation is significant at the 0.01 level (2-tailed).

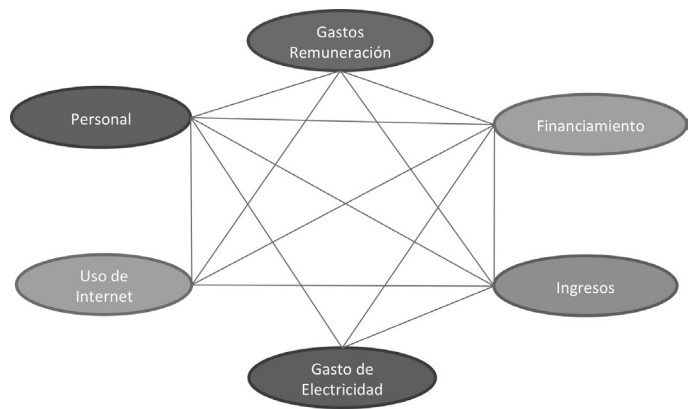
*. Correlation is significant at the 0.05 level (2-tailed).

Fuente: INEC- Censo Económico 2010

Elaboración: El Autor

Bosquejando un gráfico de las correspondiente correlaciones de cada una de las variables, tomando en consideración un valor de correlación superior a 0,05, tenemos:

Imagen 2. Iconografía de variables.



Fuente: INEC- Censo Económico 2010

Elaboración: El Autor

Debido a que existe mucha relación entre los porcentajes de las variables explicadas, es necesario obtener la explicación de cada una de las variables por otro método.

Correlaciones parciales y semiparciales

La correlación parcial es la correlación de una variable independiente (X_i) y una variable dependiente (Y) cuando

se han suprimido los efectos de las otras variables independientes tanto X_i e Y (Hair-Anderson-Tatham-Black, 2005).

Para calcular las correlaciones parciales es necesario concurrir a la siguiente fórmula:⁸

Realizando las correlaciones parciales en parejas de variables

Correlación semiparcial de Y , X_1 , dado X_2

$$= \frac{\text{Corr de } Y, X_1 - (\text{Corr de } Y, X_2 * \text{Corr de } X_1, X_2)}{\sqrt{1 - (\text{Corr de } X_1, X_2)^2}} \quad (6)$$

Para tomar como ejemplo tomemos las variables:

Y = Personal

X_1 = Gastos remuneración

X_2 = Financiamiento

Tabla 5. Resultado de correlaciones parciales de cada par de combinación de variable de estudio.

Variables estáticas	G.Re -Pers	G.Re- Internet	G.Re- Electric	G.Re- Ingres	G.Re- Financ	Pers -Ingres	Pers -Financ	Pers -Internet	Pers -Electric	Internet- Ingres	Internet- Financ	Internet- Electric	Electri- Ingres	Electri- Financ	Ingres- Financ
Ingreso	0,500	0,056	-0,92		0,259		-0,03	0,551	0,444		0,029	-0,511		-0,001	
Financia- miento	0,582	0,265	0,152	0,348		0,785		0,775	0,286	0,659		0,121	0,635		
Internet	0,675		0,303	0,424	0,433	0,645	0,354		0,418				0,806	0,467	0,570
Electric	0,675	0,444		0,528	0,796	0,871	0,526	0,838		0,830	0,518				0,606
Pers		-0,344	0,059	-0,093	0,222					0,169	0,113	-0,112	0,738	0,358	0,487
G.Rem						0,807	0,420	0,834	0,391	0,720	0,454	0,276	0,723	0,460	0,619

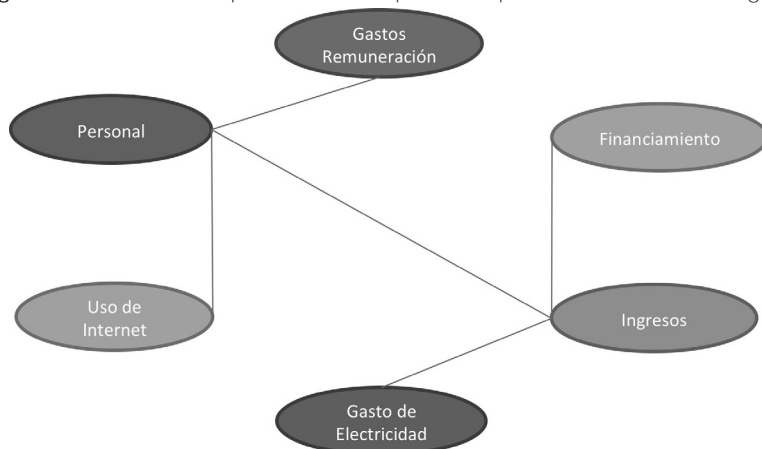
Fuente: INEC- Censo Económico 2010
 Elaboración: El Autor

Por cada una de las columnas tomamos en consideración aquellas por la cual superen el valor de 0,5. Existen solo correlaciones entre:

Gasto Remuneración –Personal Personal – ingreso Personal – Internet Gasto electricidad-Ingreso Ingreso – Financiamiento.

Esta estrategia elimina toda correlación espuria de cada par de variables de investigación en todas sus combinaciones. Se elimina cada columna por la cual no posea una correlación mayor a 0,5, para así mantener el mismo criterio de selección y eliminación.

Imagen 3. Correlaciones espurias eliminada para cada par de variables de investigación.



Fuente: INEC- Censo Económico 2010

Elaboración: El Autor

Parte 4.

Análisis de las actividades económicas

Los datos son favorables para permitir realizar un análisis mucho más profundo. Si bien es cierto que cada actividad económica ha representado una fila de datos de información, sería relevante saber qué o cuáles actividades económicas se encuentran aún más relacionadas con las variables de investigación, en otras palabras, qué variables pesan más con respecto a las actividades económicas, para ello debemos de considerar en efecto que cada fila se refleje como un instante del análisis, caracterizado por una variable indicadora igual a 1 en el ins-

tante de la línea considerada, y a 0, en otro caso: es decir, se considera lo siguiente:

Crear 20 columnas adicionales que reflejen las actividades económicas como variables.⁹

Se mostrará los valores de correlación mayores a 0,5 (Se sugiere también mostrar en los resultados las correlaciones mayores a 0,3).

Se adoptará el mismo criterio en donde se trazaran vínculos de relación desechadas por las correlaciones de Pearson.

Se bosquejará el gráfico resultante.

En la Tabla 6 se muestra cómo debe mostrarse la tabla para su análisis.

Tabla 6. Información extendida que refleja a las actividades como variables.

TOTAL	Perso- nal	Gastos remune- ración	Financia- miento	Gastos de Elec- tricidad	Gastos de Elec- tricidad	Ingresos	A 1	A 2	A 3	A 4	A 5	A 6	A 7	A 8	A 9	A 10	A 11	A 12	A 13	A 14	A 15	A 16	A 17	A 18	A 19	A 20
AA1	0,0181	0,01470	0,03569	0,00453	0,01550	0,0076	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA2	0,00670	0,11294	0,03097	0,00125	0,00215	0,0226	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA3	0,1296	0,11358	0,14923	0,09736	0,46680	0,2777	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA4	0,0042	0,00617	0,00635	0,00107	0,00358	0,0043	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA5	0,0030	0,00501	0,00117	0,00115	0,00398	0,0030	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA6	0,0159	0,01688	0,02062	0,01296	0,00732	0,0139	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA7	0,2968	0,16111	0,20889	0,30312	0,18892	0,41349	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
AA8	0,0238	0,02865	0,02604	0,02283	0,03468	0,0265	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
AA9	0,0744	0,02026	0,01675	0,03841	0,05222	0,0190	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
AA10	0,0251	0,01623	0,08779	0,12158	0,02285	0,0206	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
AA11	0,0233	0,04863	0,18050	0,01718	0,04798	0,0488	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
AA12	0,0047	0,01487	0,00900	0,01142	0,00960	0,0057	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
AA13	0,0302	0,04338	0,05258	0,10206	0,02004	0,0227	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
AA14	0,0307	0,03716	0,00744	0,03822	0,00798	0,0114	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
AA15	0,0898	0,13157	0,10965	0,01701	0,03506	0,0412	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
AA16	0,1105	0,13046	0,03525	0,07397	0,02860	0,0301	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
AA17	0,0517	0,06183	0,01183	0,07554	0,02342	0,0172	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
AA18	0,0110	0,01986	0,00126	0,01006	0,00665	0,0049	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
AA19	0,0495	0,01657	0,00886	0,04999	0,02265	0,0085	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
AA20	0,0001	0,00015	0,00014	0,00029	0,00003	0,0000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Fuente: INEC- censo Económico 2010
Elaboración: El Autor

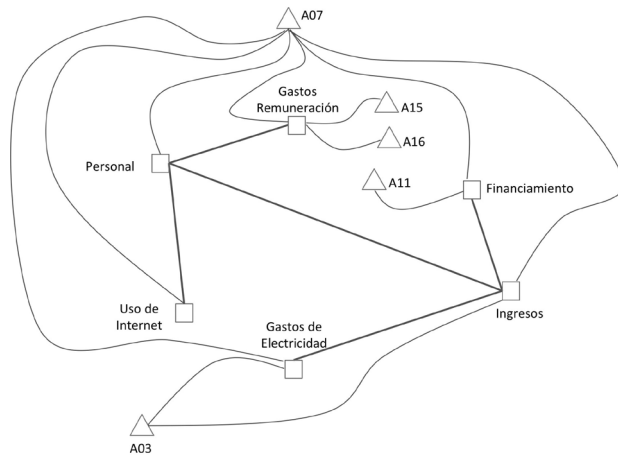
Tabla 6. Correlaciones Bivariadas de Pearson en donde se muestran las actividades económicas como variables de estudio.

	Pearson Correlation	Sig. (2-tailed)	N	Pearson Correlation	Sig. (2-tailed)	N	Pearson Correlation	Sig. (2-tailed)	N	Pearson Correlation	Sig. (2-tailed)	N	Pearson Correlation	Sig. (2-tailed)	N	Pearson Correlation	Sig. (2-tailed)	N	Pearson Correlation	Sig. (2-tailed)	N	Pearson Correlation	Sig. (2-tailed)	N
Personal																								
	-0,109	0,148	20	-0,156	0,272	20	-0,160	0,297	20	-0,116	0,434**	20	-0,163	0,246	20	-0,102	0,494	20	-0,103	0,494	20	-0,103	0,494	20
	0,648	0,534	20	0,511	0,246	20	0,500	0,297	20	0,625	0,000	20	0,604	0,000	20	0,605	0,000	20	0,604	0,000	20	0,604	0,000	20
	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
Gastos																								
	-0,165	0,294	20	-0,205	0,297	20	-0,210	0,374	20	-0,155	0,518*	20	-0,100	0,676	20	-0,158	0,507	20	-0,158	0,507	20	-0,158	0,507	20
	0,488	0,209	20	0,387	0,204	20	0,374	0,204	20	0,515	0,019	20	0,560	0,019	20	0,507	0,019	20	0,507	0,019	20	0,507	0,019	20
	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
Financiamiento																								
	-0,053	0,071	20	-0,162	0,369	20	-0,181	0,369	20	-0,109	0,590**	20	-0,124	0,507	20	-0,140	0,485*	20	-0,140	0,485*	20	-0,140	0,485*	20
	0,824	0,767	20	0,494	0,110	20	0,444	0,110	20	0,647	0,006	20	0,604	0,006	20	0,555	0,030	20	0,555	0,030	20	0,555	0,030	20
	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
Internet																								
	-0,152	0,163	20	-0,163	0,158	20	-0,163	0,158	20	-0,123	0,844**	20	-0,091	0,704	20	-0,039	0,984	20	-0,039	0,984	20	-0,039	0,984	20
	0,524	0,494	20	0,492	0,506	20	0,493	0,506	20	0,604	0,000	20	0,704	0,000	20	0,311	0,646	20	0,311	0,646	20	0,311	0,646	20
	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
Energía																								
	-0,076	0,106	20	-0,103	0,924**	20	-0,102	0,924**	20	-0,095	0,308	20	-0,034	0,887	20	-0,060	0,801	20	-0,060	0,801	20	-0,060	0,801	20
	0,749	0,656	20	0,666	0,000	20	0,669	0,000	20	0,692	0,187	20	0,887	0,000	20	0,984	0,000	20	0,984	0,000	20	0,984	0,000	20
	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
Ingresos																								
	-0,096	0,062	20	-0,103	0,514*	20	-0,106	0,514*	20	-0,081	0,821**	20	-0,053	0,824	20	-0,070	0,781	20	-0,070	0,781	20	-0,070	0,781	20
	0,688	0,796	20	0,665	0,020	20	0,656	0,020	20	0,733	0,000	20	0,824	0,000	20	0,770	0,000	20	0,770	0,000	20	0,770	0,000	20
	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20

Fuente: El autor

Realizando el gráfico correspondiente y diferenciando la actividad económica como un triángulo y la variable de investigación como un cuadrado (M, 1999).

Imagen 4. Iconografía de Correlaciones en donde se muestran las actividades económicas con variables de investigación.



Fuente: El Autor.

Conclusiones

1. Existe una correlación lineal muy fuerte (todas mayores a 0,5), eliminado inclusive toda correlación ilegítima existente entre cada combinación de variables, siendo estas correlaciones:

- “Gastos en remuneraciones totales de los establecimientos únicos y matrices censados en el 2009, según clasificación CIIU 4.0 de la actividad principal” con el “Personal ocupado, según clasificación CIIU 4.0 de la actividad económica principal a nivel nacional, 2009”.
- El “Personal ocupado, según clasificación CIIU 4.0 de la actividad económica principal a nivel nacional, 2009” con el “Uso de

la Internet en establecimientos únicos y matrices censados, según clasificación CIIU 4.0 de la actividad principal a nivel nacional, 2009”.

- El personal ocupado, según clasificación CIIU 4.0 de la actividad económica principal a nivel nacional 2009 con los ingresos anuales de los establecimientos.
- El “Monto de Financiamiento obtenido en el 2009 de establecimientos económicos únicos y matrices censados, según CIIU 4.0 de la actividad principal nacional” con los “ingresos anuales de los establecimientos”.
- Y los Ingresos anuales de los establecimientos con los “Gastos de energía eléctrica anual de los

- establecimientos únicos y matrices censados en el 2009, según CIIU de la actividad económica principal a nivel nacional”.
2. Toda variable de investigación está íntimamente correlacionada con la actividad económica “Comercio al por mayor y al por menor; reparación de vehículos automotores y motocicletas”, es decir los cambios que se puedan presentar en esta actividad económica, afectan significativamente a todas la variables investigadas en este documento.
 3. La actividad económica “Industrias manufactureras”, está muy correlacionada a los “Gastos de energía eléctrica anual de los establecimientos únicos y matrices censados en el 2009; según CIIU de la actividad económica principal a nivel nacional” y a los “Ingresos anuales de los establecimientos”.
 4. Existe también una correlación significativa entre la variable “Gastos en remuneraciones totales de los establecimientos únicos y matrices censados en el 2009, según clasi-

ficación CIIU 4.0 de la actividad principal” y la actividad económica. “Administración pública y defensa; planes de seguridad social de afiliación obligatoria” y la “enseñanza”.

5. Por último la variable Monto de financiamiento obtenido en el 2009 de establecimientos económicos únicos y matrices censados, según CIIU 4.0 de la actividad principal nacional” se encuentra muy correlacionada con la actividad económica “actividades financieras y de seguros”.

Recomendaciones

Se debería ampliar este estudio por provincias para determinar en cuáles de ellas están más relacionadas con las actividades económicas y con las variables que se investigaron en este documento; se recomienda incluso realizarlo por cantones.

Anexos

Anexo 1: Ejemplo de las obtención del Coeficiente de Correlación de Pearson

Tabla 7. Valores porcentuales de las variables personal y Gastos en Remuneración.

Actividades económicas	Personal (X)	Gastos en remuneración (Y)	(XY)
A1	1,810%	1,470%	0,00027
A2	0,670%	11,294%	0,00076
A3	12,960%	11,358%	0,01472
A4	0,422%	0,617%	0,00003
A5	0,309%	0,501%	0,00002
A6	1,595%	1,688%	0,00027
A7	29,686%	16,111%	0,04783

Actividades económicas	Personal (X)	Gastos en remuneración (Y)	(XY)
A8	2,389%	2,865%	0,00068
A9	7,443%	2,026%	0,00151
A10	2,517%	1,623%	0,00041
A11	2,338%	4,863%	0,00114
A12	0,471%	1,487%	0,00007
A13	3,026%	4,338%	0,00131
A14	3,078%	3,716%	0,00114
A15	8,985%	13,157%	0,01182
A16	11,055%	13,046%	0,01442
A17	5,178%	6,183%	0,00320
A18	1,101%	1,986%	0,00022
A19	4,956%	1,657%	0,00082
A20	0,011%	0,015%	0,00000
TOTALES	100%	100%	0,10063

Fuente: El Autor

Obtenemos los promedios de cada una de las variables;

De (3) tenemos que:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{1}{20} = 0.05$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{1}{20} = 0.05$$

y de (4) obtenemos

$$\bar{y} = \frac{\sum y_i}{n} = \frac{1}{20} = 0.05$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{1}{20} = 0.05$$

Utilizando (2)

$$Cov(x, y) = \frac{\sum x_i y_i}{n} - \bar{x} \bar{y} = \frac{0.100630}{20} - (0.05)(0.05) = 0.002531$$

$$\sigma_x = 0.06720$$

$$\sigma_x = 0.06720$$

$$\sigma_y = 0.04916$$

$$\sigma_y = 0.04916$$

Finalmente utilizando (1), obtenemos ¹⁰:

$$\rho = \frac{Cov(x, y)}{\sigma_x \sigma_y} = \frac{0.002531}{(0.06720)(0.04916)} = 0.76628$$

Notas

- 1 La covarianza mide la intensidad de variación entre dos variables
- 2 Según el archivo Presentacion.pdf de la página web del El Instituto Nacional de Estadística y Censos <www.inec.gov.ec/>
- 3 Según el archivo Metodología.pdf de la página web del El Instituto Nacional de Estadística y Censos <www.inec.gov.ec/>
- 4 Debido al largo nombre que posee la variable, se las renombrará por sugerencia del autor.
- 5 Véase <http://unstats.un.org/unsd/publication/SeriesM/seriesm_4rev3_1s.pdf> página oficial de la United Nations Division.
- 6 El documento fue proporcionado por funcionarios del INEC y entregado a las universidades del país por medio de un CD. También se encuentra en la página web <http://www.inec.gov.ec>
- 7 El documento fue proporcionado por funcionarios del INEC y entregado a las Universidades del País por medio de un CD.
- 8 Se utilizó el Programa Spss para la obtención de todas las correlaciones parciales.
- 9 Se ingresarán los datos en el programa SPSS para obtener los resultados de la Correlación Bivariada de Pearson.

- 10 Note que el valor 0,76628 es el mismo valor de la primera fila y tercera columna de la matriz de correlaciones de Pearson.

Bibliografía

- Cisepa, E. L. (2003). *El coeficiente de correlación y correlaciones espúreas*. Perú: Pontificia Universidad Católica del Perú, Departamento de Economía.
- Damodar Gujarati, D. P., *Econometría*, 5ta. ed. Mexico D.C, Mc Graw Hill, 2010.
- Hair-Anderson-Tatham-Black, *Análisis Multivariante*, 5ta. ed. Madrid, Pearson, Prentice Hall, 2005.
- Vivanco, M., *Muestreo Estadístico Diseño y Aplicaciones*, Santiago de Chile, Editorial Universitaria, 2005.
- Sitios web
<<http://www.inec.gov.ec/home/>>.
- M, L. Une nouvelle approche dans le choix des régresseurs de la régression multiple en présence d'interactions et de colinéarités. Francia: Janvier.
- Madrid, U. C. (s.f.). *Estadística Básica*. Recuperado el 1 de agosto de 2011, de <http://www.ucm.es/info/genetica/Estadistica/estadistica_basica.htm#coeficiente%20de%20correlaci%C3%B3n>, 1999.

Envío 31/agosto/2011 - Aceptación 31/octubre/2011